Training Neural Networks to distinguish craving smokers, non-craving smokers, and non-smokers

Christoph Doell[1], Sarah Donohue[2] *, Cedrik Pätz[3], and Christan Borgelt[1,3]

[1] University of Konstanz - Dept of Computer and Information Science
Universitätsstraße 10, 78457 Konstanz - Germany
christoph.doell@uni.kn
[2] Leibnitz-Institute for Neurobiology - Department of Behavioral Neurology
Brenneckestraße 6, 39118 Magdeburg - Germany
[3] University of Magdeburg - Institute for Intelligent Cooperating Systems
Universitätsplatz 2, 39106 - Germany

Abstract. In the present study, we investigate the differences in brain signals of craving smokers, non-craving smokers, and non-smokers. To this end, we use data from resting-state EEG measurements to train predictive models to distinguish these three groups. We compare the results obtained from three simple models – majority class prediction, random guessing, and naive Bayes – as well as two neural network approaches. The first of these approaches uses a channel-wise model with dense layers, the second one uses cross-channel convolution. We therefore generate a benchmark on the given data set and show that there is a significant difference in the EEG signals of smokers and non-smokers.

Keywords: Smoker · Craving · EEG · Neural Network · Classification

1 Introduction

Substance abuse and addiction have many negative effects on the health of the addicted individual, and society as a whole, with the resulting health care costs alone being staggeringly high. Understanding how addiction works in the brain is therefore of utmost importance, as it is the first step in determining the best ways to treat addiction. Nicotine is legally used worldwide and provides an excellent opportunity to study addiction in the brain for multiple reasons. First, nicotine, like other drugs, has chemical effects on the brain, which can be measured (e.g., [10]). Second, after only few hours of abstinence, smokers start to crave the next cigarette – a hallmark of addiction, the neural underpinnings of which are little understood. Third, the legality and prevalence of nicotine provides an available subject population, without the ethically and legally questionable issues that can be present when examining addiction in illegal substance abuse. Fourth, the study of addiction in humans avoids the ethically questionable administration of drugs to animal models, which may or may not respond in similar ways to the drugs as humans do.

* This work was partially funded by DFG SFB 779 TP A14N

In the present study, use neural networks to classify the data from smokers who have just smoked (non-craving), smokers who have abstained from smoking for four hours (craving), and non-smokers. The data format used, electroencephalography (EEG), measures the millisecond-by-millisecond changes in electrical potentials on the scalp, providing a measure of the neural activity over time. We used resting-state data here, in which the participants fixated on a cross for approximately 10 minutes, to determine if the patterns present in this basic data could be detected and used to classify our subject groups.

2 Related Work

Previous research on addiction has generally used functional magnetic resonance imaging (fMRI) to examine differences in resting-state data between craving and sated smokers [9], or between smokers and non-smokers [17][19], with frontal, executive-control-related regions such as the insula or dorsolateral prefrontal cortex (DLPFC) being implicated in differences present. Previous modeling techniques used machine learning to determine smoking status in fMRI data [12]. FMRI is, however, an expensive method to use, with low temporal resolution and restrictions in subject populations due to its necessary metal-free environment.

Measuring EEG Signals

EEG is a cost-effective and non-invasive technique, which can be used to assess changes in neural activity over time. The data are measured at various electrodes (in the current study, 32) relative to a reference electrode, and the electrodes cover the head in a way (see Figure 1) to optimally pick up neural signals, presumably generated from local field potentials [11]. Each electrode measures the signal representing one channel of the dataset.

Event Related Analysis

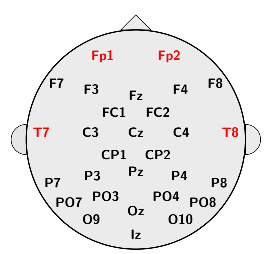

Fig. 1. EEG electrode locations.

One more traditional way to analyze EEG data is to conduct an event-related analysis. In this form, a subject is given a task, and every time an event is presented (e.g., a picture), a code signal is added to the data, which can be used for time-locked selective averaging. Using this method on a partially overlapping population with the present study, Donohue and colleagues [3] found that when smokers were craving, they showed generally more arousal in their neural activity in response to all stimuli presented, and, regardless of whether the image was nicotine-related or non-nicotine related. It is an open question, however, if overall enhanced arousal would be present in resting-state data, and if the differences observed in the event-related study are great enough to be captured by a machine learning algorithm.

Resting State Analysis

For many other big-data tasks, the important patterns are known. For example the p300 [18] shows a specific reaction to stimuli after 300 milliseconds, which has been widely used since its discovery. One difficulty of using resting-state signals is that these patterns are unknown in our case. Previous studies using EEG measures disagree on the frequency bands in which significant differences occur. For smokers, Brown [2] found reduced alpha and increased rhythmic high frequency, Rass [13] detected reduced alpha as well, but also reduced delta and Knott [6] reports reduced delta and increased beta.

EEG Analysis using Neural Networks

The recent developments in feature extraction using neural networks offers a novel way to examine brain data, to find patterns, which may be highly meaningful and would otherwise remain undiscovered. Schirrmeister [15] applied deep convolutional neural networks (CNNs) on EEG data. They share connection weights to find specific local patterns within the given data. With pictures, CNNs have proven to be very successful for object recognition on the Imagenet competition [7]. As objects in pictures are represented by groups of nearby pixels, convolutions are perfectly tailored for this task. However, it is not clear that the patterns we are looking for in the EEG are also local.

3 Data Description and Preprocessing

The experimental methods and procedures used in this study were authorized by the Ethics Committee at the Otto-von-Guericke University of Magdeburg. From all participants of the study written, informed consent was given prior the participation. All subjects were financially compensated for their time.

Initially, EEG data from 30 smokers and 9 non-smokers were measured by 32 electrodes positioned on the scalp in the frequently-used 10-20-system as depicted in Figure 1. Smokers were measured in two sessions: In the non-craving session they had recently smoked a cigarette, in the craving session they had not smoked for at least 4 hours. For non-smokers, only one session was obtained. Each recording session consists of 9.5 minutes resting state with a recording frequency of 508 Hz. Specifically, one measurement contains 508 Hz$\times$60 seconds$\times$ 9.5 minutes $\approx$ 290.000 dimensions per channel.

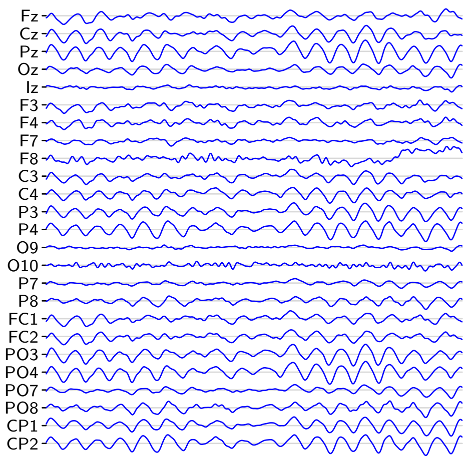

Fig. 2. Data snippet from non-smoker

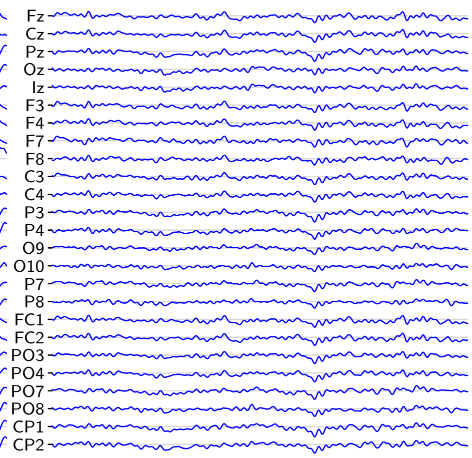

Fig. 3. Data snippet from craving smoker

Preprocessing

EEG electrodes not only measure signals arising from the brain, but they also pick up various forms of noise. As a first preprocessing step, we applied a low-pass filter at 30 Hz and a high-pass filter at 0.5 Hz. This removed high frequency noise, including power line interference, some muscle artifacts, slow-drift related movements, respiration and sweat artifacts. Subsequently, we removed physiological artifacts using Independent Component Analysis (ICA). This algorithm uses the sensor signals and creates independent components. From these, we manually selected and removed components containing eye blinks, eye movements and heart beat. The selection of the components was conservative, as the removal of a noise-like component also removes some brain signals, and it is not possible to remove only the noisy parts of a component.

We verified visually that the ICA had successfully removed these artifacts, but it had also created high frequency noise, which is why we filtered again, keeping only the signal between 0.5 Hz and 30 Hz. To additionally exclude any remaining physiological noise present in the data, we excluded channels Fp1 and Fp2 (eye artifacts) and T7 and T8 (muscle artifacts) from subsequent analysis.

We had to exclude three participants: One had fallen asleep during recording, and two more were rejected, as we could not successfully remove the artifacts without removing most of the signal as well. For the subsequent analysis, we used data from a final set of 27 smokers and 9 non-smokers, each with 25 channels. Sample snippets of approximately two seconds of preprocessed data are shown for a non-smoker in Figure 2 and for a craving smoker in Figure 3. We chose snippets that can be distinguished as easy as possible. We performed the preprocessing using the MNE framework [4].

Effects Hypothesized to be Inherent

EEG data is well known to have a bad signal-to-noise ratio, even when carefully preprocessed. We suppose that two effects could be inherent, an effect of addiction and an effect of craving. We investigate both effects by creating models for classifying two or three classes. Considering *craving smokers vs. non-smokers* should measure both effects, which, if these effects sum up, would be indicated by a high predictability. *Non-craving smokers vs. non-smokers* only measures the effect of addiction and *craving vs. non-craving* takes the effect of craving into account. The most difficult problem uses data from all three classes and tries to distinguish them all.

The measurements for craving and non-craving subjects were taken from the same subjects. This means, for training of the models containing craving and non-craving measurements, a problem with the assumption of test sets and training sets being independent and identically distributed (i.i.d) occurs if data from the same subject are used in both sets simultaneously. For a detailed description, we refer to the work of Le Boudec [8]. Although this seems like a theoretical problem, it is possible that our models find and learn person-specific patterns (i.e., identifying a specific subject) [14]. These patterns could confuse the model when a subject was in both data sets at the same time. This can cause difficulties for the model, as it gets the opportunity to learn subject-specific patterns in the EEG signals, which might be used to identify the person. To mitigate this problem, for all subjects both measurements (craving and non-craving) were used either for training or for testing. In this case it is still possible that the model learns person-specific patterns, but these will not directly affect the results.

4 Methodology

The data set contains measurements from 36 participants, which is a lot for medical studies, as measurements are expensive, but is very small for data analysis. Therefore the general reliability of the results is low, and results should be verified with more data, when available. This also motivates the need for a methodology that adds as little variance as possible.

With data from only 36 participants, a classification is prone to over-fit the training data and needs a good feature extraction, especially since our input space has $\approx$ 290.000 dimensions per channel. We focus here on neural network models, which are known for their ability to automatically detect features that are relevant for the task at hand. To reduce the problem of few measurements, bootstrapping methods exist, which generate more training examples by re-sampling the data. All samples generated from one measurement have to be considered statistically dependent on each other. This means, in order to keep the independence assumption, they may not be used in both the training and the testing set at the same time.

In our case we apply bootstrapping by taking time windows of fixed size from a measurement and use these windows instead of the whole measurement. This leads to two advantages: First, it reduces the dimensionality, second it increases the number of training samples. But it also raises questions: Which length should the window have and should windows be allowed to overlap?

A larger window length gives the model a longer signal to process and therefore more information, which might help to distinguish the classes, but it also increases the time to process the data. On the other hand a smaller window length makes it possible to generate more training samples.

Another important topic is the validation method. With few samples it is common to reuse data several times in independent tests in order to get an estimate of the quality, for an unknown, unseen data set. A good overview of cross-validation procedures was written by Arlot [1].

In the Leave-One-Out Cross-Validation *(LOO-CV)* one measurement is used for validation, while all others are used for training. For 36 subjects, the number of different splits with *LOO-CV* is only 36, which causes limited computational costs, but also lacks the possibility to perform more independent runs. This method maximizes the number of training samples but is known to return optimistic results.

The random shuffle split cross-validation copes with these problems: We split our data into training and test data at a certain arbitrarily chosen ratio. We choose 7 out of every 9 subjects for the training data and the remaining 2 for testing, i.e. 28 for training, 8 for testing in total. This means we randomly choose two non-smokers and six smokers out of the 36 subjects as test set. To minimize variance during the testing we apply stratification. This guarantees that for any random split, those ratios hold for all classes. As the numbers of smokers and non-smokers are multiples of 9 in all classes (27 craving, 27 non-craving and 9 non-smoker), no rounding is needed here. The random choice has the advantage that it can easily be repeated often to generate more reliable results. In this example, there are $\binom{27}{6}$ possibilities to choose smokers and another $\binom{9}{2}$ for non-smokers, which adds up to 10.656.360 possibilities.

With 3 times as many smokers as non-smokers, our classes are unbalanced. We handle this by balancing the class weights during the training and the validation process. To score our results, we use the class-balanced accuracy in all our experiments. Note that this score is equivalent to the class-balanced F_1-Score [16] with micro-averaging.

5 Experimentation and Results

We performed two series of experiments. In the first one we wanted to start simple. We focused on the problem to distinguish non-smokers (ns) from craving smokers (c), as we expected it to be the easiest. We looked for possibly small networks and a set of parameters that generates results better than guessing. We performed various experiments on network structures, network parameters as optimizers and number of epochs and we also varied the window length.

The second experimentation series was meant to check the other problems, to improve the results, to correct possible weaknesses and to try a different network

Table 1. Network Structures of Dense Networks

Name	Structure
Dense 1	$(25\times\ \ 5)$ merge $\times 64 \times 2$
Dense 2	(25×10) merge $\times 128 \times 64 \times 2$
Dense 3	(25×20) merge $\times 256 \times 128 \times 64 \times 2$

structure. Here, we were aiming for reliable and statistically significant results, so we needed to repeat the experiments several times.

5.1 First Series of Experiments

We found the following experiment set-up to be working. We used *LOO-CV*, non-overlapping windows of length 1000, which corresponds to pieces of approximately two seconds. Thus, we created $290.000/1.000 = 290$ training samples per measurement. Using one measurement per subject, we received 10.440 samples in 25.000 dimensions.

Our neural networks used mostly dense layers and dropout. We experimented with three different models, which contain mainly dense layers and dropout. Our smallest model is Dense 1. It starts with an independent dense layer with 5 neurons for each of the 25 channels. Their outputs are then merged into one layer of 125 neurons. Next follows a dropout (rate: .2) and another dense layer with 64 neurons. Finally, we add again dropout (rate: .1) and softmax with one neuron per class.

All three variants are summarized in Table 1. For networks Dense 2 and Dense 3 we increased the number of neurons in the layer before the merge and added further dense layers (each of them accompanied by a dropout of .2) after the merge.

Results of the First Series

Dense 1 reached 60.9%, Dense 2 59.7% and Dense 3 65,9% as average class-weighted prediction accuracy. As random guessing would achieve 50%, these results indicate that it is possible to find the combined effects of smoking and craving within EEG data.

Our analysis of the first experiment series indicated that our models have a tendency to predict the craving class. As there are three times as many craving subjects as non-smokers, this imbalance occurs as well in the test set of the *LOO-CV* and could result in overly optimistic results. Hence, in the second experiment series, we repeat these experiments with random shuffle split cross-validation. (Note: As we show detailed results for shuffle split, we omit the detailed results here.)

5.2 Second Series of Experiments

In the second series of experiments we consider all four classification problems (three 2-class problems and the 3-class problem) in order to investigate the effects of smoking and of craving separately and in combination. We use shuffle-split cross-validation with 100 repetitions in order to get unbiased and reliable results. To overcome the limited number of possible samples, we now sample random pieces with replacement permitting overlapping. In this way the number of possible pieces per measurement increases to 290.000 minus the window length. We fix the number of samples per epoch to 10.000. We also compare our results to those of the simple models: predicting the majority class, random guessing, and naive Bayes. Finally, we perform t-tests to show that our models perform significantly better than the simple models. As Schirrmeister [15] recommended, we also try a convolutional network which applies a convolution over the channels. We start with a convolutional layer with 512 filters, followed by a max-pooling by Factor 2. Then again a convolution layer with 512 filters, followed by a max-pooling by Factor 2. It follows a flatten and a dense layer with 1024 neurons. The final layer uses softmax.

Results of the Second Series

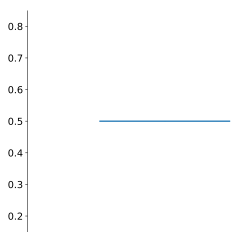

Fig. 4. Majority class

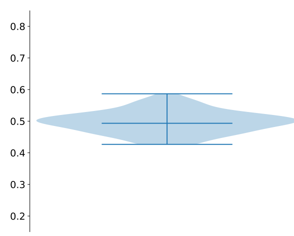

Fig. 5. Random guessing

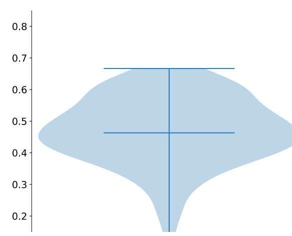

Fig. 6. Naive Bayes

For analyzing the results, we use violin plots of the average class-weighted prediction accuracy. They visualize the distribution of results and therefore show more than just mean and standard deviation. Figures 4, 5, 6 show the violin plots when predicting the majority class, when randomly guessing, and when using a naive Bayes predictor to distinguish craving and non-smokers. The first shows zero variance at a mean of 0.5. Random guessing imports some variance at the same mean value. The Bayes model has the highest variance and even a worse mean value. It is unable to detect the relevant features.

Craving vs. Non-Smoker

In contrast to the simple models, our neural networks are able to find the combined effects of smoking and craving. Figure 7 shows with 63.7% even better results than the Convolutional Network in Figure 8. Yet, the earlier 65.9% of Dense 3 were indeed optimistic.

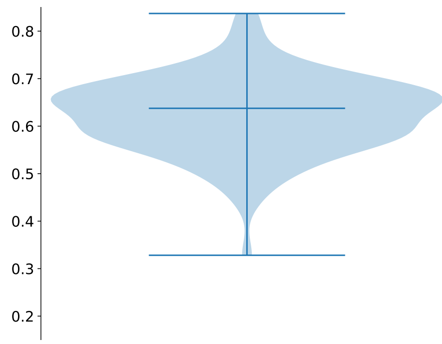

Fig. 7. C vs NS: Dense 3

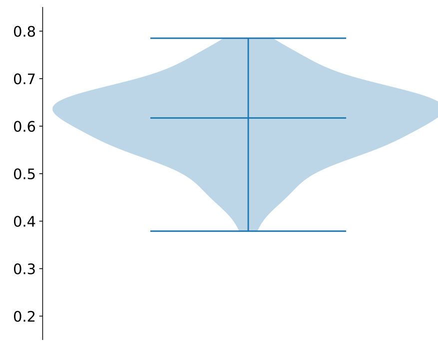

Fig. 8. C vs NS: CNN

Craving vs. Non-Craving

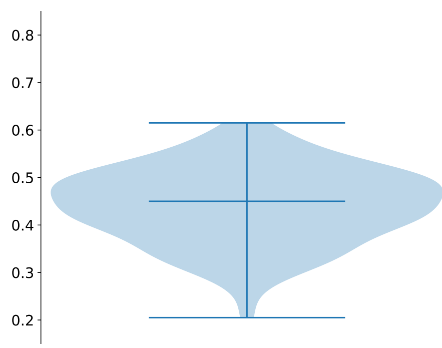

Fig. 9. C vs NC: Dense 3

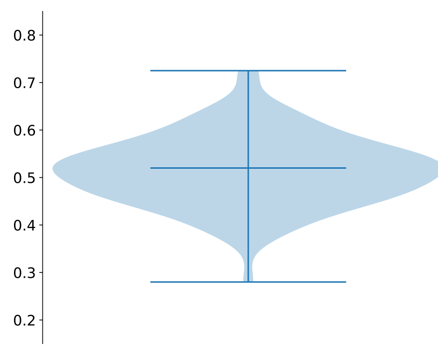

Fig. 10. C vs NC: CNN

The effect of craving in EEG data seems to be very small. In Figure 9 we see the performance of the Dense 3 network. With a median accuracy of 45% it is worse than random guessing. The convolutional network has a median accuracy of 52%. A t-test for different means comparing with random guessing returned a p-value of 0.156. This means our models are unable to predict the craving effect significantly better than guessing.

Non-craving vs. Non-smoker

The Dense model is able to detect the effect of smoking with a median accuracy of 61.8% (Figure 11) and outperforms the convolutional network (see Figure 12), which achieves only 57.8%. This shows that the effect of smoking (without craving) can be found in EEG signals.

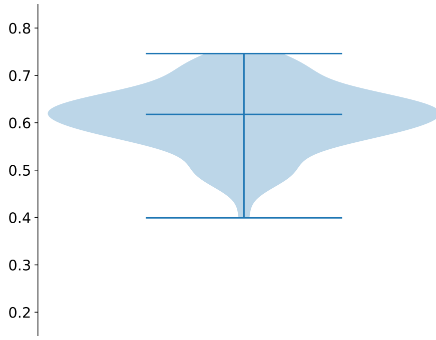

Fig. 11. NC vs NS: Dense 3

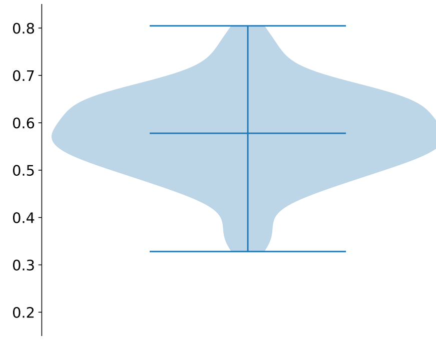

Fig. 12. NC vs NS: CNN

3-Class-Problem

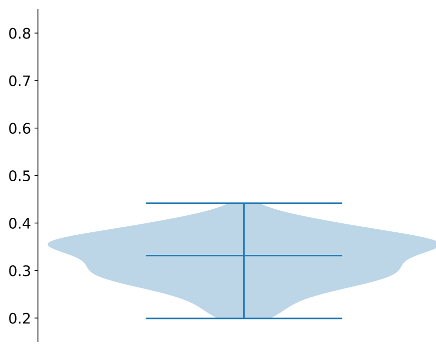

Fig. 13. 3class: Dense 3

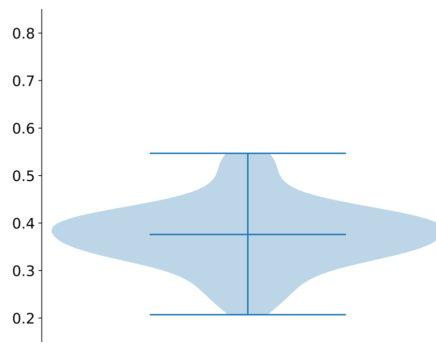

Fig. 14. 3class: Convolutional Network

For the three class problem, the convolutional network has a median accuracy of 37.6% and outperforms the dense network (Figure 13), which reaches only 33.1% – the level of random guessing. So compared to craving vs. non-smoker, the non-craving data reduced the prediction ability of Dense 3. The convolutional network shows again (Figure 14) predictions that are significantly better than guessing ($p < 0.0001$).

Finally, we look at the confusion matrix for the 3-class case. The entries contain the average normalized values and their variance as numbers and more reddish color indicates a higher mean. Figures 15 and 16 show that both models are good at correctly predicting the craving class, while both have a low quality identifying non-smokers. Also both frequently predict craving, when non-craving is correct. Therefore, this confirms that the craving effect – if existent – is difficult to find. The CNN has higher rates for all correct predictions and is clearly the better model. Nevertheless, it shows more variance in most of the cases.

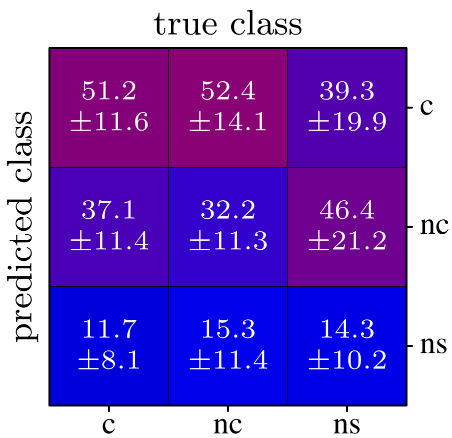

Fig. 15. Dense 3 Confusion Matrix

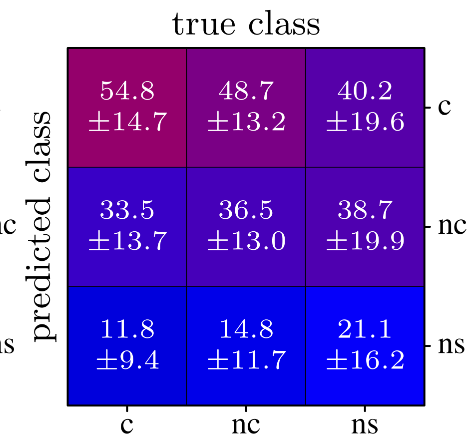

Fig. 16. CNN Confusion Matrix

6 Conclusion and Future Work

In this work, we created models to distinguish craving smokers, non-craving smokers and non-smokers. Our models are able to successfully distinguish smokers from non-smokers. Nevertheless, we found no model able to find a significant effect of craving within the data. Models distinguishing all three classes showed the same weakness.

We have shown that resting-state EEG measurements contain information on the smoking status of a person. This is a great result, especially since EEG data are known to have a low signal-to-noise ratio and thus a good classification cannot be expected. This promising finding builds the basis of future research with many implications for the study of addiction in cognitive neuroscience.

For our future work, we plan to investigate recurrent networks like the Long-Short-Term-Memory (LSTM) [5]. These units have shown good results when modeling multivariate time series, such as EEG signals. Further, we aim for visualizing the network's features in order to gain insights what are the patterns that differ in the brains of smokers and non-smokers.

References

1. Sylvain Arlot, Alain Celisse, et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79, 2010.
2. Barbara B. Brown. Some characteristic EEG differences between heavy smoker and non-smoker subjects. *Neuropsychologia*, 6(4):381–388, 1968.
3. Sarah E Donohue, Marty G Woldorff, Jens-Max Hopf, Joseph A Harris, Hans-Jochen Heinze, and Mircea A Schoenfeld. An electrophysiological dissociation of craving and stimulus-dependent attentional capture in smokers. *Cognitive, Affective, & Behavioral Neuroscience*, 16(6):1114–1126, 2016.

4. Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, et al. Meg and eeg data analysis with mne-python. *Frontiers in neuroscience*, 7, 2013.
5. Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
6. Verner Knott, Meaghan Cosgrove, Crystal Villeneuve, Derek Fisher, Anne Millar, and Judy McIntosh. EEG correlates of imagery-induced cigarette craving in male and female smokers. *Addictive Behaviors*, 33(4):616–621, 2008.
7. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
8. Jean-Yves Le Boudec. *Performance Evaluation of Computer and Communication Systems*. EPFL Press, Lausanne, 2010.
9. Caryn Lerman, Hong Gu, James Loughead, Kosha Ruparel, Yihong Yang, and Elliot A Stein. Large-scale brain network coupling predicts acute nicotine abstinence effects on craving and cognitive function. *JAMA psychiatry*, 71(5):523–530, 2014.
10. HNA Logemann, KBE Böcker, PKH Deschamps, C Kemner, and JL Kenemans. The effect of the augmentation of cholinergic neurotransmission by nicotine on eeg indices of visuospatial attention. *Behavioural brain research*, 260:67–73, 2014.
11. Steven J Luck. *An Introduction to the Event-Related Potential Technique (Cognitive Neuroscience)*. A Bradford Book, 2005.
12. Vani Pariyadath, Elliot A Stein, and Thomas J Ross. Machine learning classification of resting state functional connectivity predicts smoking status. *Frontiers in human neuroscience*, 8:425, 2014.
13. Olga Rass, Woo Young Ahn, and Brian F. O'Donnell. Resting-state EEG, impulsiveness, and personality in daily and nondaily smokers. *Clinical Neurophysiology*, 127(1):409–418, 2016.
14. Vitaly Schetinin, Livija Jakaite, Ndifreke Nyah, Dusica Novakovic, and Wojtek Krzanowski. Feature extraction with gmdh-type neural networks for eeg-based person identification. *International journal of neural systems*, page 1750064, 2017.
15. Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human eeg. *arXiv preprint arXiv:1703.05051*, 2017.
16. Fabrizio Sebastiani. Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1):1–47, 2002.
17. Luke E Stoeckel, Xiaoqian J Chai, Jiahe Zhang, Susan Whitfield-Gabrieli, and A Eden Evins. Lower gray matter density and functional connectivity in the anterior insula in smokers compared with never smokers. *Addiction biology*, 21(4):972–981, 2016.
18. Samuel Sutton, Margery Braren, Joseph Zubin, and ER John. Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188, 1965.
19. Barbara J Weiland, Amithrupa Sabbineni, Vince D Calhoun, Robert C Welsh, and Kent E Hutchison. Reduced executive and default network functional connectivity in cigarette smokers. *Human brain mapping*, 36(3):872–882, 2015.